SCOTTISH INSTITUTE FOR RESEARCH IN ECONOMICS

SIRE DISCUSSION PAPER

SIRE-DP-2015-31

Values, Multiculturalism and Representations

Martin K Jones

UNIVERSITY OF DUNDEE

www.sire.ac.uk

Values, Multiculturalism and Representations

Martin K Jones*

Abstract: The theory of values is underdeveloped within economics with very few theoretical models explaining how values are created and propagated through a population. Those models that do exist are limited in how well they explain the role of values and how they behave. This paper argues that the theory of values needs to be fundamentally rethought, incorporating basic ideas from psychology, anthropology and philosophy that have largely been ignored in the literature. The resulting model is applied to answer some fundamental issues surrounding the notion of multiculturalism.

JEL Classification:

Keywords: Values, Representations, Mental Actions, Multiculturalism

*Economic Studies, School of Social Sciences, University of Dundee, Dundee. DD1 4HN.

M.K.Jones@dundee.ac.uk tel: +44 1382 385164 fax: +44 1382 384691

1. Introduction

The influence of culture- and in particular cultural values- is an active and promising area of research within economics, resulting in a variety of field and experimental studies in the literature (See de Jong (2009) for an overview). It is becoming increasingly obvious that values have a huge impact on economic and other behaviour. Different values within a society can lead to political disagreements and clashes, different business values lead to different employment practices, different consumer values lead to different goods being bought and different values between societies can have a serious effect on the ability to trade and invest in different countries.

However, the theoretical literature has lagged behind with few papers actively trying to model values as distinct from norms and conventions. Within cultural economics, values are generally instrumentalised from questionnaires that are in widespread use in anthropology and sociology (See de Jong (2009) and Inglehart et al. (1998) for examples) and have little theoretical foundation. This is a shame in that a reasonable theory of values would help to explain some of the culture clashes that do occur and allow us to predict which situations are likely to occur. Furthermore, the papers that do exist do not seem to fit the notion of values very well in that they ignore certain common characteristics that should be explained by any reasonable theory. It will be argued that to explain these characteristics requires us to take seriously concepts that are widely used in cognitive disciplines but are not widely used in economics.

In order to demonstrate the applicability of these ideas we will construct a model of values that casts some light on the issue of multiculturalism and the issues brought up by multiculturalism. Values (together with cultural objects, icons etc.) constitute a culture (see Sperber 1996). This means that a theoretical study of values may have something to tell us about how likely particular cultures are to survive in competition with each other. Is a multicultural society likely to survive or are some cultures going to be assimilated by others? How crucial are institutions within society in the preservation of multiculturalism? The simple model presented here will give some preliminary answers to these questions.

2.Other Theories of Value

Values are quite often ignored in economics even when they are the ostensible subject of discussion. An example of this is in the book "Economics, Values and Organization" (1998) edited by Ben-ner and Putterman. In this book are papers by a wide variety of academics from various disciplines. What is interesting is that among the economists (e.g. Robert Sugden, Ken Binmore, Timur Kuran) values, as such, are only mentioned in passing or they are conflated with other moral concepts such as norms. An example is in Sugden's paper: in spite of a long discussion (p.74-77) on the importance of values, Sugden ultimately identifies values with conventions and normative expectations. However it is not clear, as I will discuss below, that values on one hand and norms and conventions on the other are in any way the same thing.

Two attempts that *have* been made to theorise values have been by Tabellini (2008) and Bisin and Verdier (2001). In both papers it is assumed that altruistic traits are passed down through a process of "imperfect empathy" from one's parents. Parents are "imperfectly altruistic" in that they judge the best interests of their children using their own, rather than their children's, utility functions. Parents then expend effort in trying to socialise their children with these cultural traits through an expenditure of effort. Children sometimes pick up these traits but at other times pick up a trait from the population as a whole. Children only definitively pick up a trait when their parents and the population as a whole share the same preferences. Tabellini extends this model to account for the fact that some individuals may be more socially or physically distant than others.

As it stands, this analysis is attractive in that it describes much about the transmission of cultural traits within a population. In particular, it neatly describes the problems of parents trying to instil particular values in an environment that may be hostile to such values and also the idea that one tends to value people further away less than those close to. However, there is little in either paper to explain how values emerge within a population. Why should one set of values be seen as something valuable that should be passed on from one generation to another? In other words these papers, while insightful, say little about how values emerge or come to be held extensively within a population in the first place.

A deeper problem with these papers is that it is not clear what a value actually is. It is described as a trait that is inherited to a greater or lesser extent from parents (or role models) to children which in turn gives utility to the parents. However, it is unclear why these traits are valuable to parents or why they think that these traits are valuable to their children. Traits are merely things that are transmitted with greater or lesser effect from one generation to another. They are obviously preferred in some sense to other possible traits. However these traits do not seem to be preferences in themselves. Furthermore, it is assumed that self - interest is separate from values but this is not obviously true- some values may be based on self-interested motivations (one may value being in a trade union because of the benefits it brings to oneself.)

Furthermore, socialisation is something that is very loosely defined. It is obvious that parents would need to make a value judgement in order to socialise their children. It is also an implication of these models that the children take on these values in opposition to their own preferences, which implies that socialisation is to some extent coercive. However, this does not seem to be the case in many circumstances where socialisation, whether from parents or wider society, seems to be voluntary. It seems that people are quite willing to become socialised into another culture and "fit in" with the local population. However, the details of this are left vague by the extant models and need to be examined in more detail.

3.Multiculturalism

In this paper the model that will be put forward will be used to give a preliminary, simple analysis of multiculturalism. There are many different views of multiculturalism so from the point of view of this paper we will focus on the ideas put forward by one author. The author selected is Parekh (2006) who emphasises the role of culture as a system of beliefs and practices where values play a central role. Everyone in Parekh's view belongs to a cultural community which shares both practices and beliefs. Society and culture are to be distinguished from each other since society consists of a group of humans and the relations between them while culture provides the "software"- the content and principles underlying those relations.

Culture, to a large extent, shapes and influences individuals. In particular, the shared beliefs of people within the cultural community mean that people are more likely to follow a given practice. However, culture is malleable so that people within a culture are free to leave the culture (albeit retaining some vestiges of it even after having rejected it.). Cultures, and the individuals within those cultures, are able to take on ideas and values from other cultures. Diversity is inevitable (and, according to Parekh, desirable) within societies and it is quite possible to have cultures living side by side with each other within a society. Even within an initially homogenous society, the diverse capacities of humans within the society mean that there will develop considerable divergences from the cultural community's "norm".

This view of multiculturalism has become quite influential within cultural studies so it is worth asking whether the implicit claims made hold up to scrutiny. Is it indeed possible for cultures to survive in a stable state next to each other in a society? If humans are indeed free to look at and take on board values from other cultural communities will this lead to the cultures surviving or will one set of values simply collapse? Do divergences from cultural norms necessarily remain or will there be a tendency for these divergences to vanish (or even take over the entire cultural community?).

4.Grounding for a theory of values

The aim of this paper is to present a theory of values that explains the formation and distribution of values within a society. As such, the model is an attempt to formulate how values may form in a "state of nature" without any institutions[1] and with a population that is free to learn from all other members of the population. It also excludes any influence from education or from parental guidance. This means that members of the population formulate their ideas about values and their actions given these values without any indoctrination or coercive socialisation[2] of values.

By its very nature, these restrictions exclude a large proportion of methods by which people do come to hold their values. As a matter of fact, most people probably do get their values from one of the sources mentioned above. One example is in education. Many people are taught fundamental values while being educated (objectivity in scientific research for example). These values do not necessarily arise from interactions between individuals in a population. Indeed, the values may be formulated by a comparatively small group of people (say, scientists and philosophers throughout history) while the education system is used to spread these values to the wider population.

The model also ignores the effect of role models in a society. Fashion, for example, is an example of a set of values that tends to be set by a comparatively small group of people (actors, models, socialites, pop stars) and is spread through the wider population by the mass media. In medieval times local holy men may have played the same role for religious people.

[1] See Denzau and North (1994) for an example of a theory where institutions *do* influence values.

[2] It will be argued later that "non-coercive" socialisation is little different from ordinary learning

Role models have many more links to other people when it comes to persuading people to change their mind about values. It is assumed in our model that such people do not exist.

The model, therefore, is narrowly defined. However, its aim is to abstract away from the complications outlined above and find out how values can be defined and modelled in a simple model. Once it is seen that values can be modelled with clarity and precision then it will be easier to incorporate the complications.

I will assume throughout this paper that humans are capable of non- self- interested behaviour and motivations. The paper will take no position on how humans arrive in this position in evolutionary terms. It will be assumed that non- self- interested behaviour has minimal implications for the *fitness*[3] of the individual involved and so fitness will not enter into the utility functions of the individuals making altruistic choices. Individuals will take notice of payoffs involved insofar as these impact on utility but the impact may not be self-interested. One may donate to charity, for example, and gain greater utility from donating more money.

5.Values as Preferences

First, it is necessary to define values. One definition is by Rokeach (1973):

"To say that a person 'has a value' is to say that he has an enduring belief that a specific mode of conduct or end-of-state existence is personally and socially preferable to alternative modes of conduct or end-states of existence"

It can be seen that this definition is very broad. Indeed one might sensibly ask what the difference is between values on one hand and preferences as defined by economists on the other. It will be argued in this paper that there is no real difference: values are simply a type of preference (although there are preferences, a sudden craving for ice-cream for example, that are not values).

One crucial distinction between values and other types of preference are their comparable longevity. One can say that someone has a value when they hold a definite preference for something for a long time. It would be peculiar to say, for example, that one valued one's wristwatch if one thinks that one will keep it now but was thinking about throwing it away five minutes ago.

[3] Indeed it is hard to see why fitness is emphasised by so many game theorists when discussing humans. One can carry out many altruistic actions without this in any way reducing one's ability to reproduce. A quick examination of demography shows that there is no positive correlation between reproductive success and wealth.

One can also say that values are consciously and intentionally held. They are not the result of accidental selection or unconscious urgings. They may be held as habits without being deliberately chosen every time they are invoked but this does not mean that they are unexamined. Even habits usually start out as conscious decisions and are usually the result of intentional behaviour.

A crucial aspect of values (contrary to Sugden's claim mentioned above) is that they are distinct from norms and conventions. A convention is a mode of action that is repeated over a long period of time as the result of an implicit agreement between the parties carrying out the action. This is usually modelled in evolutionary game theory as a particular Nash equilibrium in a game with multiple equilibria. If the population converges on one of the equilibria then that equilibrium becomes the convention. A norm is a convention that evolved previously which is upheld even when there is no history of interaction between the agents who are currently interacting. One uses a normative action because one has legitimate expectations (Sugden 2005) that the other person will also use that norm.

However, a critical aspect of values is that they may hold even if there is no action played corresponding to that value. To take an example: one may hold that stealing kestrel eggs from nests is wrong without ever seeing a kestrel egg that was in a position to be stolen. Likewise, one may value patriotism in wartime without ever having to fight a war. In both cases there is no underlying convention that forces one to follow the moral rule. Instead, one follows the rules because of one's underlying values.

It follows from this that values cannot necessarily be accessed by revealed preference. Revealed preference would reveal the existence of conventions and norms but would not say anything about values that are not exemplified by actions. This is reflected in the practical sphere by the use of questionnaires to elicit values rather than revealed preference techniques (c.f. Inglehart et al. 1998).

Values, as defined Rokeach, can cover both individual and group values. While individual values are important to the person concerned, economics is more concerned with the actions of large groups of people rather than the idiosyncrasies of individuals or small groups. As such, this paper will focus on the behaviour of values held by large groups of people rather than individuals.

Another important point is that values seem to be context sensitive (Seligman & Katz 1996). Different values seem to be invoked in different circumstances. When different issues are at stake, different values are invoked to argue them out. A ready wit, for example, is more valued at a dinner party rather than at a funeral wake. This has interesting implications:

values are not as universally or consistently applied as one might think and one might apply values differently in different circumstances. It follows that any theory of values must take account of this context sensitivity.

Another point that has to be made about values is that they seem to vary consistently between groups. These may vary between groups in a society (e.g. different food restrictions in the Sikh and Muslim religions in India) or between societies (Inglehart 1998). The latter is particularly important in cultural economics where these differences between countries drive many of the differences in economic variables that have been found in the literature (de Jong 2009). These differences in values have been tested experimentally (see Henrich et al. 2004) and there are distinct differences in reactions between different groups.

The experiments by Henrich et al also demonstrate that any model of values must allow for the fact that different groups may end up with different values in the same situation. It would be unreasonable to postulate, for example, that there is a "long-run equilibrium" set of values to which all people in a given situation will converge. Furthermore, one has to account for the fact that when different groups undertake different actions they seem to give different reasons for these actions as well. It is not just that people converge to different equilibria within the same framework. Often their mental frameworks seem to differ as well.

6. Fundamental ideas for a theory of values

Discussion of values is widespread across a variety of disciplines. In this section we will discuss the ideas of two writers- one a philosopher, Joseph Raz (2003), and the other an anthropologist, Dan Sperber (1996). These two writers have similar views on many aspects of values and have tried to accommodate many of the points outlined above, albeit from differing perspectives.

Raz approaches the problem of values from a philosophical perspective. His aim is to try to provide a view of values that allows for value pluralism without degenerating into relativism. In his view, values initially emerge through the use of supporting practices. As people carry out these practices then values are created by the participants in these practices to regulate and excel in them. An example of this is opera which is a practice that has a variety of values (good operatic singing, intelligent interpretation of the music) that could only exist together within opera.

These values, once they are established through the practice, have an independent existence of their own. They can survive even if the original supporting practice dies out. These values tend to be highly specific in that they only apply in the tightly defined areas where they emerge. However, once they emerge then they can be used anywhere within this area without needing a supporting practice. More general values emerge as generalisations of more specific values and do not have their own sustaining practices.

It should be pointed out that Raz is not claiming that sustaining practices "create" values- valuation is a human activity that can only be done by human beings. Humans value things but which things are to be valued is dependent on the social practices in existence. Raz therefore is not claiming that valuation has any objective basis but merely that the occurrence of values is based on contingent practices.

Sperber (1996) approaches culture from an anthropological point of view. In his view the basic building block of culture (which includes values) is the *representation*. The term "representation" has a wide range of meanings but for, the purposes of this paper, the type of representation we (and Sperber) are interested in is the *mental* representation[4]. A mental representation is an interpretation of the world held internally by an information processing device (such as humans!). This is a *cognitive* picture of how the human mind works- relating the mind to information- processing by a computer.

This notion of the mind as being modelled as a computer has its origins in the work of Simon (Simon 1959). Under Simon's ideas, humans are seen as boundedly rational computing machines who construct "specifications" (Simon & March p. 172) of the situations in which they operate. These "specifications", which can be seen to be equivalent to the "representations" used by Sperber and others, consist of knowledge of future events, knowledge of alternatives available for action, knowledge of consequences attached to alternatives and rules or preferences for ordering consequences. This defines a situation for a particular actor but does not give an objective assessment. Indeed, this assessment is usually a screened, simplified and biased view of the situation. Knowledge of the future for example is quite often simply unknown and is usually dealt with by using previously used ideas.

This idea of representations is used widely within Social Psychology. According to Ross and Nisbett (1991 –see also Bowles & Gintis 2011 p.43) situational variables have a substantial effect on human decision making and quite often make more of a contribution towards decision making than the character and stable preferences of the decision maker. Furthermore the representations of situations are construed by the decision maker and so are essentially subjective rather than objective. They can vary from person to person and can vary within a person from one time to another. Indeed, it has been argued by Sen (1980) that

---

[4] A similar notion of representations as the grounding for values is put forward by Mandler (1993- see also Ross and Nisbett 1991), although he tends to refer to a particular type of representation known as a *schema*.

description of any kind is inevitably selective and partial so insofar as a representation is an individual's description of the external world, it will suffer from the same flaws for the same reasons.

The lack of knowledge of the situation and the arbitrariness involved is particularly marked when it comes to values because of the lack of any objective link between values and a particular objective situation (See Bowles & Gintis 2011 p. 44 for a similar claim). As Raz points out, valuation is essentially a human activity and any attempt to derive values from objective reality is probably futile. It follows that any representation of a situation when a person is valuing that situation will not have totally objective foundations. No situation will logically suggest any particular set of values independently of human valuation. Representations therefore have an inescapably subjective element when referring to values.

Mental representations, according to Sperber, are used by the information processing devices (i.e. minds) to affect other people by interfacing with the physical environment. This can be by direct or indirect communication and this in turn causes other people to create mental representations in their own minds. As a result of this, mental representations are transmitted from one person to another. Sperber then defines a *cultural or social representation* as a representation that is commonly held among the general population. He sees the aim of anthropology as being to investigate the "epidemiology" of these cultural representations.

One question that Sperber attempts to answer is which sort of mental representation will succeed in becoming a social representation? This is a wide ranging problem but some relevant attributes are:

i) Ease of memorisation

ii) Existence of relevant background knowledge.

iii) A motive for communicating the content of the representation

iv) Recurrence of the situation which the representation gives rise to.

v) Availability of external memory stores (writing, mobile phones)

vi) Existence of institutions engaged in transmission of institution.

This will tend to cut down the number of mental representations that can ever become social representations. One side effect of this is that many representations that are deemed to be important in society- knowledge of the law or of science for example, rarely become common knowledge among the mass of the population because of their complexity, lack of knowledge etc.

7. Interpretation and variation of representations

Both Sperber and Raz emphasise the fact that one's knowledge about values is not fixed. According to Raz, one's knowledge about values is limited and it is quite possible that in the same situation, two people may quite reasonably disagree. Raz locates this disagreement in the human capacity for judgement and understanding. Given limited information, people will use their judgement to interpret a situation and come to a conclusion. These judgements may be starkly different and this leaves scope for considerable disagreement between people[5].

Sperber also emphasises the role of interpretation, although his ideas are somewhat more radical than those of Raz. In his view, there are two points of interpretation. The first point is when representations are first conceived. This leads to various distinct interpretations of a situation. The second point when interpretation happens is every time a mental representation passes from one person to another. Communication therefore is as much about transformation as it is about transmission. It should be noted that Sperber does not claim that transformation results in the emergence of something completely different. Indeed, some resemblance to the previous representation remains.

Sperber creates a model of cultural attraction in which descendants of a particular mental representation, when communicated, always differ from their parents. However, this variation is not simply random. Certain versions of mental representations are more "attractive" than other mental representations. As a result of this, representations will not transform at random but will tend to converge on these representational attractors.

There are various factors which make a particular version of a mental representation more attractive than others. Sperber theorises that the crucial factor in human cognitive processing is the maximisation of relevance i.e. having a maximum effect for minimum effort. Some things that are perceived as relevant are highly idiosyncratic and only apply to a few people. However, some are generally held by large numbers of people. These must be easy to remember and to understand. There must also be enough incentives to recall the representation and transmit it. Also it must be credible- it must be believed by the people who transmit it and it must also make sense. It would be expected therefore that any representation must contain beliefs that are easily understood, that make sense or are worth having.

[5] Ross and Nisbett (1991) come to a similar conclusion from a social psychological perspective based on the work of Asche (1940)

Sperber argued his case for representation in opposition to the common idea that cultural progress is carried out by evolutionary style selection arguments (c.f. Cavalli-Sforza & Feldman 1981, also Dawkins 1976). However, it has been argued by Henrich and Boyd (2002) that this contrast is a false one. In a series of three models, they demonstrate that the variation in representations caused by such strong attractors can be modelled with selection processes and ultimately, even if the selection process is comparatively weak, it will come to dominate the overall evolutionary process.

It can be said therefore that representations are essentially subjective entities that can vary considerably from person to person. However, while some of these representations may "mutate" as they spread through the population, these mutations are rarely stable and there is a tendency for them to converge on a prototype. These mutations can be modelled as a random process of deviations from a mean where the latter corresponds to Sperber's attractor representation.

8.Values and Emotions

A crucial assumption made in this paper is that self-interest is not the only motivating force for human beings (while still being *a* motivating force). This should not be a controversial assumption as a multitude of economics and psychological experiments have shown that, even in economically important decisions, self-interest is not always the deciding factor (c.f. Bowles and Gintis 2011, Fehr 1999). As has been pointed out by Mansbridge (1998), it is impossible to seriously base morals on self-interest since it goes against the innate emotions and cognitive capacities of human beings.

Another problem with the self-interested view is that *one can make an argument for self-interest being a type of value*. A cynic might argue that free-market ideologues do this all the time but there are a vast number of situations where being able to follow your own self-interest is generally valuable. Simply choosing which apple a person would like to eat is a privilege that most people enjoy and find valuable yet it is fundamentally self-interested. Self-interest therefore is not really a special state apart but rather an alternative set of values that a person may select from those that are available.

Mandler (1993) emphasises the crucial role played by values in the inducing of emotions. Mandler points out that, however visceral and "hot" an emotion may be, all emotions must rely on some cognitive processing so that a person may establish and interpret a situation. A person usually only gets angry at having their wallet stolen if they have, indeed, had their wallet stolen. A person who gets angry at having their wallet stolen when this is not the case is seen as having no justification for their anger. To get angry therefore requires cognitive processing to establish that the wallet has indeed been stolen.

Many, but not all, emotions tend to be aroused by an expectation generated by a mental representation. This of course does not apply to all emotions. Many emotions are instinctual and automatic so that there is little beyond identifying the situation that the mind does in sparking the emotion. If an iron rod is being dropped from the roof of a building then, once this has been identified, the body tries to avoid it, automatically invoking a feeling of fear. However, many such emotions are learned and are acquired over time and it is these emotions that often emerge from following, or failing to follow, values. People, therefore, will identify the situation and also the value that needs to be followed before emotion emerges in response. Given that values are assumed to be preferences expressed within a certain representation, it follows that many people will feel emotional if values are being violated or even if values are being followed.

Many emotions therefore have their foundations in values or in the violation of perceived values. Frank (1988) has observed that emotions can override self-interest and that this enables people to make commitments which solve problems that cannot otherwise be solved[6]. However, the ability to make commitments means that it must be possible to reliably communicate one's emotions. Frank claims that this is achieved either through a person building a reputation for honestly having these emotions or by outward signs of this emotion that are very hard to fake[7].

It should be noticed that Frank is not claiming that these emotions are on public view or that they cannot be faked. However, Frank points out that that, like all signalling devices, emotions are quite hard to fake and also that, for emotions to work as a commitment device, most emotions must be genuine. While it may be difficult to find out another's true emotions, it is rarely impossible and fakery is often unsuccessful. Indeed, people who tend to be successful at faking their emotions (such as psychopaths) tend to be ostracised within society and quite often end up worse off. Indeed, from an evolutionary point of view, it is difficult to see exactly why emotions would survive if they were easily faked. Their role as a signal would be undermined by rampant fraud, while being psychologically costly to those experiencing the emotions.

[6] An example of this would be the emotion of love in marriage which would prevent a person from bailing out at the earliest possible favourable opportunity. A person who can show that they love another person would be seen as a better marriage partner because they are in the marriage for the long term and so they can invest more in the marriage.

[7] See Scharlemann, Eckel et al. (2001) for an economic experiment that demonstrates that participants react positively to photographs of their opponents smiling in a trust game.

It follows that we have a linkage between values and emotion with the causality going from the values to the emotions rather than vice versa. The existence of values induces certain kinds of emotion when they are followed or not. Emotions are very hard to fake and are also observable by other individuals. Since they are hard to fake then they act as a reasonably reliable signal to the other individuals as to the underlying values being expressed.

9. Building blocks of a theory of value

The discussion above has shown that there is room for a theory of value which incorporates the notion of representation from cognitive science. This paper will follow through this idea in incorporating it within a game-theoretic model. However, this creates problems because there is no accepted way in which representations can be modelled within game theory. The usual method of looking at heterogeneity among agents is to use the concept of type. However, this is obviously insufficient since types are assumed to be fixed for each person and out of the conscious control of the player (Harsanyi 1967) while one can always change one's representation.

Another method that has been used to allow for context in games is framing (see Bacharach 2006 on variable frame theory). However, framing is essentially a property of the external world while mental representations exist in the mind. The contrast between the two can be seen in an advertising slogan "Beeples washes whiter than other low-cost brands!". This frames Beeples as an efficient clothes detergent. However, this may not affect how people perceive Beeples as a brand i.e. their mental representation. (Beeples may have a history of causing skin complaints for example). As we have discussed earlier, we will focus on the non-objective view of representations (albeit one that is widely held within the population).

In order to motivate the following discussion we shall introduce the notion of a *base game*. A base game is a normal form game structure $G=\langle P,S,Q\rangle$ where P is the set of n players, S is an n-tuple of pure strategy sets (one for each player) and Q is an n-tuple of outcomes. A base game, it should be noted, has no explicitly defined payoffs but only outcomes specifying the event that occurs. An outcome may have some material payoffs or it may have something which happens which a players likes or approves of. However, this is not incorporated within the structure of the base game. Instead the base game illustrates the number of players and the strategies given to each player, while leaving the payoff structure open. In this paper we will assume that we are dealing with two player- two strategy base games.

Formally, a representation can be seen as a possible allocation of utilities to the base game. There will almost certainly be more than one possible allocation in much the same way as there will be more than one possible representation. This means that different representations will allocate different utilities to the base game. The intuition behind this is that representations are different according to different contexts and different players as described in section 6. As a result of these different perceptions, greater or lesser preference will be given to one outcome in a game matrix over another. Representations therefore will reflect the context-sensitive nature of values in terms of utilities that vary according to the situation and the representation of that situation.

How representations come into being is still the subject of intensive research. However, representations can be summoned from memory, they can be copied from another, similar, situation by a process of analogy (c.f. Gentner et al 2001), they could be acquired as part of religious instruction or parental guidance or they could be picked up from the mass media or by malicious rumours. There are a multitude of ways in which representations can be acquired and many of them can conflict with each other. We will assume, following Sperber, that these representations are *social* representations.

Furthermore, more than one representation can be held in the mind at one time. This does not mean that one has to subscribe to all the possible representations that one has in one's mind. If a person is a voter in a United Kingdom General Election then one may, for example, subscribe to Labour party values and vote for the Labour party while having a perfectly clear idea of Conservative party values. Likewise, it is possible to change one's mind between different representations. There are two ways in which this could happen. One is that a representation may change so that additional justifications for a representation make it seem plausible. Another is that circumstances change so that a formerly rejected representation seems to be a more plausible way of perceiving the world. In this paper we will focus on the mechanics of the latter.

According to Davidson (1963), actions are caused by reasons. Representations, since they consist of a mix of beliefs and desires, provide these reasons in the form of belief-desire pairs- each action being justified by one of these pairs. We will assume that individuals are *minimally rational* in the sense that, within a given representation, the reasons for actions are not contradictory. So, for example, one cannot choose an apple because one likes the taste of apples but reject a pear because one thinks all fruit tastes disgusting. However, it is quite likely that, when comparing reasons for actions across representations, the reasons in different representations will contradict each other.

Assuming minimal rationality leads to a conclusion about modelling- we cannot in general model a person as having two active representations at once i.e. individuals cannot play mixed actions across two different representations as this would require the individual to entertain contradictory reasons underlying the actions. For this reason, any modelling will have to be in a situation where individuals only have one active representation at once. This suggests that, in population based models, one would model values using, for example, evolutionary models where movement comes through individuals changing from wholly subscribing to one representation to wholly subscribing to another.

As has been argued above, this paper rejects the idea that agents are *necessarily* self-interested. Instead, self-interest is one possible attitude that one could incorporate into a representation. Such a representation would have utilities that are monotonic with the material payoffs in the outcomes. As such, the representation would compete on an equal basis with other possible representations. There is no presumption here in favour of self-interest although it is possible that in many situations, such as when one is trying to preserve one's life, the utilities in the self-interested representation would be very high.

The rejection of the idea of self-interest as the *only* motivational component of a representation is what allows the framework put forward here to work. If self-interest is imposed as necessary then the scope for differences between representations would be cut down to those that are monotonic with the material outcomes- which would be a very meagre offering.

Given the multiple possible motivations, self-interest is modelled within this paper as being just one amongst many possible motivations. In fact no distinction is made between self-interest and other motivations in that there is no separation between self- interested preferences and non-self-interested preferences. All are subsumed under generic preferences which may be self –interested or otherwise. Whether a person behaves in a self-interested manner depends on the representation. It is likewise assumed that all preferences, whether self-interested or not, can be incorporated into a single utility scale. None of this explicitly contradicts the von Neumann- Morgenstern theory of utility and games which simply requires that preferences satisfy the axioms of expected utility.

Strictly it will be assumed that the utilities assigned by representations will be *expected* utilities comprising probabilities and utilities. The utilities represent the valuations prescribed by the representation. The probabilities represent the belief that certain things may happen or are the case. This encapsulates the idea that culture and values are partially determined by the beliefs of the cultural community. These beliefs may be associated with, say, religion or maybe the importance of certain historical events. A Scotsman may, for example, value wearing a kilt because he has a belief that it is an item of clothing that has a history over a thousand years old[8] and also attaches a high utility to that historical background. These probabilities are *not* connected to the proportions of people with particular representations in the population.

A 2x2 base game can therefore be transformed into an ordinary 2x2 game by the use of a representation. Assume that this is an evolutionary symmetric game being played by one population of players. In that case the payoff matrix will be as follows:

	Player 2		
Player 1		Up	Down
	Up	a,a	b,c
	Down	c,b	d,d

Where a,b,c,d are all utility payoffs. In that case, according to Weibull (1995) there are four categories of game:

Category 1:a>c; b>d; this leads to a Nash Equilibrium at (Down, Down)

Category 2: a>c; b<d: this leads to two Nash Equilibria at (Up,Up) and (Down, Down)

Category 3: a<c; b>d: this leads to a mixed Nash Equilibrium at $(\frac{b-d}{(c-a)+(b-d)}, \frac{b-d}{(c-a)+(b-d)})$.

Category 4:a<c; b<d; this leads to a Nash Equilibrium at (Down, Down)

In this paper it will be claimed that the preferences under given representations attached to the base games are to be identified with values.

Definition: A value is a preference relation between outcomes within a particular representation.

[8] This is a false belief- see Trevor-Roper (1983). Beliefs need not be true to be included in a representation. All such probabilities are assumed to *subjective*.

Even though the model has not been fully explained, one can see that this method has the advantage of satisfying some of the criteria for a theory of value. First of all, not all values can be detected by a researcher through revealed preference. In a category 2 representation, the population may converge to playing (Down,Down) but this simply reveals that d>b. It does not reveal that a>c. The latter would have to be discovered by a researcher by other means. This fits in with the idea that values do not always correspond to actual choices.

Secondly, the use of representations is context sensitive. Different contexts evoke different representations and these in turn evoke different utilities as payoffs in the game. The use of representations therefore explains why a given base game with the same actions (e.g. whether or not to tell a joke) may result in different actions being chosen in different contexts (a dinner party versus a funeral). Representations also explain why different cultures choose different actions in similar choice situations. This is because different cultures provide different reasons and hence different representations for choosing one action over another.

However, it is insufficient simply to assert the existence of representations. One has to understand why agents within a society come to hold one representation rather than another. In particular, since economists and other social scientists are interested in humans *en masse* rather than individuals, why some representations are more widespread than others. One also has to understand why such representations are long- lived- why do they survive while others die out comparatively quickly?

To start to answer these questions one needs to understand how representations become attached to base games. One can discern three possibilities. The first is where there is an already existing *unique* representation that can be retrieved from memory or can be acquired by copying other people or by comparing with analogous situations. In this case one simply applies the given representation to the base game and actions are carried out based on the utilities provided. The second circumstance is where there are no representations for a base game at all. This would occur in a novel situation where the agent has no knowledge about how to proceed. In such a situation the agent must deliberate about what to do in this situation and create a new representation, possibly based on analogies with partially similar situations and partly from applying general principles. Once created such a representation then becomes available for other situations in a similar way to Raz's theory of values.

The final case is where there are two or more representations available for a given base game. This case is more complex and will be the focus of the rest of the paper. It will be assumed that these representations all pass the criteria laid out by Sperber for a social representation i.e. they are easy to remember and understand, there must be some reason for holding them and they must be credible. If this is the case then the representations will in likelihood be held by large numbers of people.

This leaves us with the problem of how agents pick a representation of a particular situation to assign utilities to the base game. It is claimed here that the process of picking a representation is a conscious intentional process by the agent. As such it is a species of *mental action*. Mental actions are processes in the mind that have as their goal another mental process (Proust 2001). Examples of mental actions are easy to come across. One may try to remember when the bus is going to leave. A purchaser may try to work out what the VAT is on a purchase of window frames. A student tries to concentrate on a lecture. In each case an agent is deliberately trying to engage some mental faculty.

This notion of mental action is quite old and goes back to the Enlightenment. Locke (1689/1998) believed that the mind deliberating and operating on itself and its own ideas was the bedrock of philosophy. Geach (1957) presents a more recent 20th Century view of mental acts, particularly in relation to judgment. Mental actions are therefore much the same as ordinary, physical actions in that they have a goal, are intentional and are not the result of unconscious thought. As intentional actions they can be rationalised by the beliefs and desires attached to the action- in other words the beliefs and desires motivate the action and act as a causal explanation for that action (Davidson 1963).

It follows that one can model the picking of one representation out of several[9] as a simple choice between actions as one would do with ordinary physical actions. This way of modelling representations is not unusual as Henrich and Boyd's (2002) first model is essentially a model of this process within a population. However, the model presented here is the first to put this explicitly within a game- theoretic context.

Each representation is motivated by beliefs and desires and the usual way to represent beliefs and desires in economics is through the use of probabilities and utilities. This means that not only does each representation allocate utilities to the outcomes of a base game but also there are utilities attached to the choice of each representation. It is assumed that utilities attached to outcomes can be easily integrated with the utility attached to choosing a representation.

One possible problem that could arise is that, when one is assigning a set of utilities to a base game, the utilities are qualitatively different across representations. This is because a

[9] Note that this notion of choice of representations is similar to a choice in descriptions as outlined by Sen (1980)

utility under a representation is conditioned on that representation and hence that particular way of understanding the situation. It will be assumed, without further investigation, that these utilities can be made consistent with each other across representations[10]. It should be noted that this is simply a measurement issue related to the consistency of the expected utility measures- it does not bear on the fact that different representations may have *reasons* that contradict each other.

Given that a choice is made between representations then the question has to be asked as to why representations cannot be changed simply to suit the convenience of the agent. If one ends up in an unfavourable situation as a result of adopting a representation then why can one not simply rewrite one's representation? It will be seen that the choice of representation is part of an extended game where one is brought into equilibrium in strategies in the base game but also in representations. Representations are therefore fixed by being in equilibrium. Rewriting a representation creates an additional strategy in a new game which also has its own equilibrium which also fixes the representations.

As mentioned above, we will focus here on changes in choices as a result of changes in numbers in a population holding a particular representation and the utilities attached to that representation. As Ross and Nisbett (1991) point out, other people are generally one's best source of information on the world around us and so the initial spread of representations in the population will assumed to be by a process of diffusion of social representations. It has been shown that most people dislike being in tension with the group of people around them and will not only vary their behaviour but also their *understanding* of a situation in order to fit in. This suggests that not only will one's strategy choice in the base game change in line with that of the group and one's utilities but also one's choice of representation as well.

It is assumed, following the discussion in section 8 that values evoke emotions and that these emotions can, fallibly, be observed by other people. It follows that we will make the same assumption in the model which we are building and we will assume that the other player's chosen representation will be available to a given player. While it is acknowledged, following Frank (1988) that this is costly, this costliness will not have any effect on the model since the costs will be the same whichever representation is chosen. To avoid the complications caused by some members of the population not knowing another's representation, we will simply assume that everyone knows which representation is chosen by one's opponent.

[10] One possibility is to impose contraction and expansion consistency on the representation- dependent utilities as in Sen (1997). However, this goes beyond the scope of the paper.

10. The model

The preceding discussion leads on to a very simple modelling strategy for values. Every person who plays a game first of all chooses a representation and, given the utilities attached to the base game, then chooses an action to play in the base game. We will refer to the choice of (behavioural) strategy at the representational level as a choice of representation while the choice of behavioural strategy in the modified base game will be called the *physical strategy*.

Choices of physical strategy in the modified base game and representational choice is determined simultaneously in equilibrium. In this paper we will concentrate on the simplest case where there are two available representations and two strategies for each of the two players. Given that each player has two representations, this suggests that there are four different combinations of representation possible.

The model has the structure of a sequential, two-player, symmetric, evolutionary game with one population and no moves by nature. More precisely, it is a two stage- two strategy, simultaneity game where payoffs are cumulative between stages (Cressman 2003 p. 192). This means that, when playing a game the two players simultaneously choose their representation and then, once the representations have been chosen, they choose which action to play given the utilities established by the representation. This leads to an extensive form game with four subgames, each in the form of a 2x2 normal form game. It is assumed that the player knows which combination of representations is held by both players before they choose their actions. A stochastic element will be added to the model to represent variation in utilities.

Cressman (2003) has put forward a method by which extensive form games can be modelled by evolutionary methods while still taking into account the sequential nature of actions within the game. Cressman demonstrates that one can decompose a replicator dynamic for the normal form of the whole game into separate replicator dynamics for the subgames and the truncated game. This can be done by assuming that the replicator operates on the Wright Manifold- the combination of probability points where the choice of strategy in one subgame is independent of the choice of the same strategy in another subgame. The game as a whole can be brought into equilibrium by first of all finding the equilibria of the subgames and then substituting one from each into the truncated game. It can be shown that when the subgame and truncated game replicators converge on an equilibrium point then this is a subgame perfect equilibrium for the whole game.

The game as a whole is played by one population. In normal form this is fairly straightforward as this just requires a one-population set of equations for the strategies played in the game. When the game is decomposed according to the Wright manifold then, for the truncated game and any symmetric subgame, the situation is still the same. Each one is treated as a smaller, one-population game. In the case of a two-stage, two strategy simultaneity game this would mean that each symmetric subgame and the overall truncated game would have a one-equation replicator dynamic equation.

However in such symmetric extensive form games, there will always be asymmetric subgames as well as symmetric subgames. In this particular type of game there will be two such games. This means that it is not appropriate to treat these games as symmetric. However, it is also not appropriate to treat the two sides of the game as members of different populations as it is assumed that the whole game is played with one population. In this case Cressman recommends treating the two games as components of a single linked asymmetric game where each "player" in the game is actually a role selected within the game by a player from a unified population. One set of choices leads to the player playing one side of the game while another set of choices leads to them playing the other side of the game.

The replicator dynamics used in this model are used as an approximation to social learning rather than having a biological context. As such one can see members of a population as having a tendency to play a given strategy, the greater the difference in expected utilities between one strategy and the average expected utility of all strategies. The biological interpretations of the replicator dynamics involving concepts such as fitness, reproduction and multiple biological generations are ignored here (see Binmore 1988 for one possible interpretation along these lines).

Each game played by the population has the structure in figure 1. R_1 and R_2 represent the representational choices made by player 1 and player 2. The information set covering player 2's nodes implies that each person chooses their information set simultaneously, although the nature of their choices is known to each other later on in the game. The subgames starting from u_1, u_2, u_3, u_4 represent the different interpreted versions of the base game. The subgames emerging from u_1 and u_4 represent the situation where both players have chosen the same representation. The asymmetric subgames following from u_2 and u_3 represent situations where the two players have chosen different representations.

{Figure 1 Here}

The strategies in the game correspond to the two types of action discussed above. R_1 and R_2 correspond to the mental actions of choosing representation 1 and representation 2. U and D are physical actions corresponding to moves in the base game. These are the moves that actually result in a material outcome being achieved.

It is assumed that any utility deriving purely from the use of representations is incorporated into the final utilities in the terminal nodes. It is also assumed that, if a player chooses a certain representation and physical action (given their opponent's physical actions) then they will have the same payoff irrespective of the representation chosen by their opponent. This means that a player's final payoffs are not influenced by the representation chosen by one's opponent.

The use of this model allows us to model many of the basic facets of the discussion above. The choice of representation is a mental action where individuals deliberately choose which representation to use in analysing the situation. The extensive form structure of the game allows this mental action to be taken prior to the physical action and represents a settling of the mind of the player on one particular interpretation. The simultaneous choice of representations by agents models the fact that agents do not know each other's choices before they make their own.

It is assumed that players know each other's representations before they play their physical actions. This reflects the previous discussion of Frank's and Mandel's work where it was assumed that one's view of the world would be accessible to the other player as a result of the visibility of their value-induced emotions. Meanwhile the evolutionary structure allows one to explain the intuition common to Raz and Sperber that values are essentially social phenomena that emerge through social interaction in the population.

More specifically, the model also catches Raz's intuition that the occurrence of values depend on social practices. If one sees the physical games being played as social practices then the representations that are selected out by the evolutionary processes can be seen to depend for their existence on those practices[11].

[1111] It may be wondered why the agent doesn't simply pick a representation according to his own self-interest. However, this model allows for other attitudes apart from self-interest- among them an interest in accuracy and objectivity which would pull against self-interested motives. On the other hand, this model does allow for the possibility that self-interest may overcome an interest in accuracy. This would be the case where the agent suffers from wishful thinking.

Similarly there is a stochastic element in the model to model Sperber's idea that representations are continuously reinterpreted. The "attraction" model put forward by Sperber suggests that agents can interpret situations in a variety of ways that lead to deviations from the original "attractor" interpretation. However these deviations eventually converge on the attractor. Since interpretations of a situation cannot be predicted, this reinterpretation process (following Henrich & Boyd 2002) is best rendered by a stochastic process in which the mean of the process stands in as the attractor.

It is further assumed that the main effect of this reinterpretation process is to change the utilities in each outcome and does not affect choices of physical strategies. As it is usually assumed in game theory that everything affecting preference would be included within the utility numbers, this is followed here. It should be noted (following Sperber 1996) that this idea of stochastic choice is population based in that different interpretations diffuse through the population. This means that there is no sense that the reinterpretations cancel each other out. Any other effects of reinterpretation are assumed to be irrelevant and are not included.

This notion of stochastic preferences is related to the idea of random utility (Becker 1963). It is assumed that each payoff is a random utility consisting of a baseline (or mean) utility with a stochastic variation. Under each representation there is a different stochastic variation although it is assumed that within each representation the stochastic terms are the same.

Note that the game allows for two representations to be chosen which lead to four subgames (with roots at u_1,u_2,u_3 and u_4). The subgames following from u_1 and u_4 will be referred to as *symmetric* subgames and correspond with the following normal form matrices (ignoring the stochastic element and looking at the mean utilities):

	Player 2		
Player 1		U	D
	U	a,a	b,c
	D	c,b	d,d

And

	Player 2		
Player 1		U	D
	U	w,w	x,y
	D	y,x	z,z

It follows that there are two asymmetric subgames following from u_2 and u_3. Given the use of Cressman's model these can be seen as two different roles in one conventional asymmetric game as shown in the matrix below:

	Player 2 (From u_3)		
Player 1 (From u_2)		U	D
	U	a,w	b,y
	D	c,x	d,z

This means that, when a player reaches u_2, then they will take the role of player 1 in a two-population asymmetric evolutionary game while if a player reaches u_3 then they take the role of player 2 in that game.

The structure of the games means that the subgames act as representations of the original base game. Nothing is said within the model as to the relationship between any material payoffs in the base game and the representational subgames within the game in figure 1. The model does not deny that some links may exist but simply does not make them explicit. This is completely conventional- strictly, within Von Neumann and Morgenstern's (1953) game theory, *all* payoffs are in terms of utilities and none of the material payoffs are specified. In addition to this, the conventional theory of utility makes no assumptions about self-interest and neither does the current theory. This means that the scope of the theory is very broad.

Furthermore, the current theory can be seen as an extension of that of von Neumann and Morgenstern. While their theory deals with a state where only one representation is available to people in a population, the current theory allows for the much broader idea that there are multiple representations, each having a different distribution of utilities. Likewise, the notion of mental action is implicit in game theory as an attempt to focus on the game and to make a decision based on the utilities (which have implicitly been assigned by a unique representation). It is only with the introduction of multiple representations that the role of mental actions becomes explicit as a choice between representations.

For the purposes of this paper the representations consist of symmetric games using the four categories outlined in section 9. It is assumed that even if the two representations in each game are in the same category (i.e. if the utilities in each representation have the same ordering) they still have different cardinal utilities. This gives some idea as to what happens when one has similar but not identical representations of a situation or when different parts of a population have different levels of "enthusiasm" for a similar view of a situation.

In order to abbreviate the representational content of different models the following notation will be used:

The term (XvY) will be used to denote when a game involves a representation formed from category X and a representation formed from category Y. So, for example, (1v2) means that the game has two representations from categories 1 and 2. The representation from category 1 would be represented by utilities a, b, c,d while the representation from category 2 would be represented by utilities w,x,y,z. Given the overall symmetry of the game, it can be seen that there is no difference if the two categories are interchanged.

11. Model Layout

Given the modelling strategy used above, then the extensive form game shown in figure 1 can be modelled as a one population evolutionary game using the replicator dynamic. We will assume that that, given a mean utility of u, the standard deviation of the utility would be σ_{R1} under representation 1 and σ_{R2} under representation 2 where R1 and R2 stand for "Representation 1" and "Representation 2" respectively. It is assumed that the stochastic terms varies over time according to the Wiener process.

First of all, examine the normal form of the game outlined above. This game is one with eight strategies (see appendix). The proportions of the population playing a given strategy s_i is q_i. p represents the mixed strategy played by one's opponent, while the payoff from playing a strategy s_i is represented by π (s_i/p). The stochastic differential equation for playing s_1 is given by:

$$dq_1 = q_1 \left[(\pi(s_1/p)dt + \sigma_{R1}dW_1) - \left[\left(\sum_{i=1}^{4} q_i(\pi(s_i/p)dt + \sigma_{R1}dW_1) \right) + \left(\sum_{i=5}^{8} q_i(\pi(s_i/p)dt + \sigma_{R2}dW_2) \right) \right] \right]$$

Where σ_{R1} and σ_{R2} are standard deviations and W_1 and W_2 are Random variables from a Wiener process.

Similar equations can be constructed for q_2, q_3 etc. These equations, as is shown in appendix 1, can be decomposed into one stochastic equation for the truncated game and four deterministic equations for the subgames. Suppose that P_{R1} represents the probability that R_1 is chosen in the truncated game while $q_U^{u_x}$ is the probability that U is chosen in the subgame following u_x (x=1,2,3,4).

The stochastic equation for the truncated game is:

Eq (1)……

$$dP_{R1} = P_{R1}(1 - P_{R1})\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4)\big)dt + P_{R1}(1 - P_{R1})\sigma dW.$$

Here, $\bar{n}_x$ are the equilibrium expected values of the subgames following from u_x .

Also:

$$\sigma = \sqrt{(\sigma_{R1}^2 + \sigma_{R2}^2)}$$

And

$$W = \frac{(\sigma_{R1}W_1 - \sigma_{R2}W_2)}{\sigma}$$

It should be noted that this looks similar to the equation in Fudenberg and Harris (1992) and indeed has many of the same properties in that negative population shares are avoided and the boundary points are steady states of the system, never achieved in finite time. However, the process by which this equation is created differs substantially from that outlined by Fudenberg and Harris. In Fudenberg and Harris the stochastic shock is attached to the system when it is expressed in terms of absolute population sizes using particular strategies. Specifically, it is added to the utility term in the population based replicator dynamic. Fudenberg and Harris then transform this stochastic replicator equation using Ito's

Lemma on the *proportion* of a population using a particular strategy. They justify this procedure by claiming that, under the biological approach, the absolute population is more fundamental and it is easier to interpret stochastic shocks as due to aggregate effects.

However, the framework adopted here uses a *social* interpretation of the replicator dynamic and, given that the source of the stochastic process is the result of aggregate "interpretational shocks" i.e. social causes rather than natural causes, it makes little sense to interpret the replicator dynamic in terms of absolute population levels.

The four deterministic equations for the subgames are as follows:

Equation 2

$$\dot{q}_U^{u_1} = P_{R1} q_U^{u_1} \Big[q_U^{u_1} a + \big(1 - q_U^{u_1}\big) b - \big[q_U^{u_1} \big(q_U^{u_1} a + \big(1 - q_U^{u_1}\big) b \big) + (1 - q_U^{u_1})(q_U^{u_1} c + \big(1 - q_U^{u_1}\big) d \big] \Big]$$

Equation 3

$$\dot{q}_U^{u_2} = (1 - P_{R1}) q_U^{u_2} \Big[q_U^{u_3} a + \big(1 - q_U^{u_3}\big) b - \big[q_U^{u_2} \big(q_U^{u_3} a + \big(1 - q_U^{u_3}\big) b \big) + (1 - q_U^{u_2})(q_U^{u_3} c + \big(1 - q_U^{u_3}\big) d \big] \Big]$$

Equation 4

$$\dot{q}_U^{u_3} = P_{R1} q_U^{u_3} \Big[q_U^{u_2} w + \big(1 - q_U^{u_2}\big) x - \big[q_U^{u_3} \big(q_U^{u_2} w + \big(1 - q_U^{u_2}\big) x \big) + (1 - q_U^{u_3})(q_U^{u_2} y + \big(1 - q_U^{u_2}\big) z \big] \Big]$$

Equation 5

$$\dot{q}_U^{u_4} = (1 - P_{R1}) q_U^{u_4} \Big[q_U^{u_4} w + \big(1 - q_U^{u_4}\big) x - \big[q_U^{u_4} \big(q_U^{u_4} w + \big(1 - q_U^{u_4}\big) x \big) + (1 - q_U^{u_4})(q_U^{u_4} y + \big(1 - q_U^{u_4}\big) z \big] \Big]$$

The decomposition of the normal form equations and the elimination of the stochastic elements is demonstrated in Appendix 1.

It will be noted that all of the equations have the form of replicator dynamics up to the multiplying of the formula by either P_{R1} or $(1-P_{R1})$ which has no effect on the direction of the trajectories. Note that equations (3) and (4) are interlinked which reflects their role as describing two roles in one interlinked asymmetric game.

Given that these are now, effectively, replicator dynamics over two- strategy games the equations can be rearranged as follows:

Equation 2*

$$\dot{q}_U^{u_1} = P_{R1} q_U^{u_1}(1 - q_U^{u_1})\left[q_U^{u_1}(a - c) + \left(1 - q_U^{u_1}\right)(b - d)\right]$$

Equation 3*

$$\dot{q}_U^{u_2} = (1 - P_{R1}) q_U^{u_2}(1 - q_U^{u_2})\left[q_U^{u_3}(a - c) + \left(1 - q_U^{u_3}\right)(b - d)\right]$$

Equation 4*

$$\dot{q}_U^{u_3} = P_{R1} q_U^{u_3}(1 - q_U^{u_3})\left[q_U^{u_2}(w - y) + \left(1 - q_U^{u_2}\right)(x - z)\right]$$

Equation 5*

$$\dot{q}_U^{u_4} = (1 - P_{R1}) q_U^{u_4}(1 - q_U^{u_4})\left[q_U^{u_4}(w - y) + \left(1 - q_U^{u_4}\right)(x - z)\right]$$

The stochastic equation for the truncated game needs further analysis to demonstrate what happens asymptotically to the proportions of representations in the population:

Define three integrals:

$$I_1 = \int_0^{P_{R1}(0)} exp\left[-\int_\alpha^{P_{R1}} \left[\frac{2\theta(P_{R1})}{\beta(P_{R1})^2}\right] dP_{R1}\right] dP_{R1}$$

$$I_2 = \int_{P_{R1}(0)}^{1} exp\left[-\int_\alpha^{P_{R1}} \left[\frac{2\theta(P_{R1})}{\beta(P_{R1})^2}\right] dP_{R1}\right] dP_{R1}$$

$$M(P_{R1}) = \frac{2}{\sigma^2} exp\left[\int_\alpha^{P_{R1}} \left[\frac{2\theta(P_{R1})}{\beta(P_{R1})^2}\right] dP_{R1}\right]$$

Where $\theta(P_{R1})= P_{R1}(1 - P_{R1})\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4)\big)$

And: $\beta(P_{R1})= P_{R1}(1 - P_{R1})\sigma$

$P_{R1}(0)$ is the initial position.

From these we can use a revised version of a proposition from Fudenberg and Harris that will establish the asymptotic behaviour of the truncated game equation.

Proposition 1 (Revised from Fudenberg and Harris 1992):

1) If $\bar{n}_2 - \bar{n}_4 > \frac{\sigma^2}{2}$ and $\bar{n}_3 - \bar{n}_1 < \frac{\sigma^2}{2}$ then $P_{R1} \to 1$ as $t \to \infty$ with probability 1
2) If $\bar{n}_2 - \bar{n}_4 < \frac{\sigma^2}{2}$ and $\bar{n}_3 - \bar{n}_1 > \frac{\sigma^2}{2}$ then $P_{R1} \to 0$ as $t \to \infty$ with probability 1

3) If $\bar{n}_2 - \bar{n}_4 > \frac{\sigma^2}{2}$ and $\bar{n}_3 - \bar{n}_1 > \frac{\sigma^2}{2}$ then $P_{R1} \to 1$ as $t \to \infty$ with probability $I_1/(I_1+I_2)$ and $P_{R1} \to 0$ as $t \to \infty$ with probability $I_2/(I_1+I_2)$
4) If $\bar{n}_2 - \bar{n}_4 < \frac{\sigma^2}{2}$ and $\bar{n}_3 - \bar{n}_1 < \frac{\sigma^2}{2}$ then

$$P(\lim_{t\to\infty} inf P_{R1} = 0) = P(\lim_{t\to\infty} sup P_{R1} = 1) = 1$$

Also the system possesses a unique ergodic distribution $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$ to which the distribution of P_{R1} converges as $t \to \infty$.

This proposition establishes the long run behaviour for categories 1 and 4 in representations demonstrating that as long as the deterministic process is strong enough then the population will converge on representation 1 for category 1 payoffs and on representation 2 for category 4 payoffs. It also shows that category 2 converges on one of the two pure equilibria representations depending on which basin of attraction the initial state starts in. In the category 3 case we know that the system converges on an ergodic equilibrium distribution.

One issue that must be discussed before analysing the model is the issue of stability within the model as a whole. In order for equilibrium to emerge for the model as a whole, Cressman (p. 193-4 Theorem 7.2.1 and footnote 4) points out that there cannot be cycles within the subgames. This is not a problem for the symmetric subgames where the decomposed replicator dynamics act as if within one population games and so all Nash equilibria are strict. However, within the asymmetric subgames this is a problem since they act like a linked two population game where cycles are possible.

The main problem is the so- called "Buyer Seller" game which is a non-degenerate, two- strategy bimatrix game (see Cressman p.78). In this case the replicator dynamics trace out cycles and do not converge towards the unique mixed Nash equilibrium. An examination of the (2v3) and (3v2) category combinations applied to the asymmetric subgames shows that the corresponding linked asymmetric game is identical to the Buyer-Seller game. Since no other category combinations in the linked asymmetric subgames are identical to the Buyer-Seller game and the Buyer-Seller game is the only one that causes these cycles it follows that the (2v3) and (3v2) category combinations are the only ones that cause this problem.

It follows, therefore, that the (2v3) and (3v2) category combinations should be excluded from any analysis of the model if one wishes to achieve a stable equilibrium. However, this makes perfect sense within the literature, especially in Raz's (2003) theory of values. For Raz, values emerge because of the existence of a sustaining practice. For something to be a practice it follows that it must be stable. Given that, under this model, a strategy played in a subgame is a practice then it can only be stable if it is in a stable, non-cyclical equilibrium. It follows that (2v3) and (3v2) must be excluded on theoretical grounds as well as mathematical grounds.

Another theoretical implication of the exclusion of the (2v3) and (3v2) category combinations is that there are no stable mixed equilibria in the asymmetric subgames. This is because the asymmetric subgames are parts of an asymmetric linked game where the two roles in the game act in the same way as two different populations. In such a case (2v3) and (3v2) are the only combinations that result in a unique interior equilibrium. In all other cases the interior equilibrium either doesn't exist or is not stable.

12. Specific properties of the model

The model outlined above is designed to explain the incidence of particular values within a population. Each representation encompasses a different set of utility numbers to be attached to the base game. Each utility number is a value in that it shows, given a particular representation, how much one outcome is valued. Another representation would give another set of values in a game. The model given here shows us how different sets of values may compete and survive against each other.

As such, we would expect this model to tell us something about culture in general insofar as culture consists of value judgements. How far can different value- systems coexist within a society when the society satisfies the model assumptions? Will one value system be wiped out by the dominance of the other one or is there a natural tendency for value systems to survive quite happily together? Such questions have obvious relevance for multicultural societies where different value systems may exist within the same society.

This leads on to particular questions that can be asked within the model. Naturally, this model is limited in that it only focusses on two possible representations and two possible physical strategies. Any real world system could have large numbers of both. However, the model does allow us to think clearly about how physical actions (i.e. the observable actions) interact with representations (i.e. people's values).

One point that must be made straightaway is that, within the model, an overwhelming majority of the possible permutations of representations end up with the entire population holding one representation. This can easily be seen by examining the properties of the truncated game. The truncated game is essentially a one-population, two- strategy evolutionary game where the strategies are mental actions selecting representations. In most cases we end up with the four possibilities outlined by Weibull (1995) for generic games which are the same as the four categories used for subgames[12]. For three possibilities the population will end up with just one representation. It is only for the mixed equilibrium (Category 3 applied to the truncated game) that we will end up with more than one representation within a population. It follows that it is rare to have a situation where more than one set of values exists in equilibrium within a population.

Another point is that the vast bulk of the possible permutations of representations also end up with the entire population playing one strategy. This can be seen by realising that for the population to end up playing different strategies involves either mixed strategies in the truncated game or in the subgames (or both). Since only one case in the four possibilities outlined by Weibull (1995) allows mixed strategies (together with the case outlined in proposition 2 below) whether in the truncated game or the subgames then it follows that the majority of the time the population will follow just one strategy.

It follows that, in the bulk of cases, there will be little evidence of differences between different parts of the population. However, this does not mean that there will be no interesting cases. One question that is of interest is whether it is possible to have a stable situation where one part of the population has a set of values that results in them playing a given physical strategy while the other part of the population has a different set of values that results in them playing the second physical strategy. This speaks to an important part of multicultural ideas as to whether different cultures can coexist in the same population.

Proposition 2:

If a population converges to a situation where two representations coexist in a population in a stochastic equilibrium governed by the probability distribution: $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$ then a player with one representation plays just one pure physical strategy and a player with

[12] It is possible (such as (1v1) or (4v4)) to end up with repeated payoffs in the truncated game. However, this simply restricts the equilibrium possibilities only to pure strategies and will result in convergence to a unique representation.

the other representation plays the other physical strategy iff the category combinations are (4v1) or (1v4).

Proof: See Appendix 1

It follows that such a situation is possible- one *can* have uniform blocs where everyone with one set of values does one thing and everyone with another set of values does another. However, theoretically, this is a rare situation compared to the likelihood of the population converging on one set of values or there existing a mixture of values and physical strategies. Furthermore, perhaps ironically in a multicultural context, this situation can only occur where the values are diametrically opposite to each other.

Another question that is of importance is whether we can have different values within a population but with everyone playing the same physical strategy? This is of obvious importance in a multicultural society as the different values would not "matter" and potential clashes would be avoided. One would have different routes to the same conclusion. Unfortunately this possibility is not allowed within the current model:

Proposition 3:

It is not possible to have a situation where there is a stochastic equilibrium probability distribution $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$ over proportions of representations held by a population and where all physical strategies are the same pure strategy.

Proof: See Appendix 1

Even allowing for this, it is interesting to see whether we can have the opposite situation. In other words, can we have a situation where we have two sets of values coexisting and both strategies are being played by people holding either set of values? In such a situation we would have a "totally mixed" population in which representations and physical strategies can be played in any combination. Again, this would reflect a valued situation within a multicultural society.

Proposition 4

A population will have a stable equilibrium in representations with probability distribution $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$ and with a mixture of physical strategies played under both representations only if the game has a category combination (3v3).

Proof: See Appendix 1

Apart from this, the playing of mixed strategies tends to be more where a population converges on one representation. This, again, is a result of the operation of the replicator dynamic on the truncated game. In most preference permutations the population will simply converge to a pure strategies equilibrium. In such a situation, the playing of multiple strategies within a population will be the result of playing a category 3 game in either or both of the symmetric subgames. Ironically, therefore, differences in values would not be a major force in driving behaviour and differing plays of physical strategies would most often emerge from everyone having the same set of values.

13. Discussion

The model outlined above is a simple construction aimed at creating a basic framework in which we can discuss the origins and spread of values. As such it ignores many of the basic channels through which values are spread such as via parental nurturing, education, the mass media, law etc. The aim has been to strip the creation and spread of values down to its absolute basic framework. However, the model is different from many of the biology- influenced models put forward before. Values in this model are not seen as unexplained “traits” passed down the generations or across a population but rather as a special type of preference which is characterised as being stable, context sensitive and widely held. The aim in this model is to take account of the psychological underpinnings of values in terms of emotion, interpretation and contextual sensitivity.

Given this, a simple evolutionary model was created that incorporated these issues into an extensive form game. This allows values to be split apart from physical actions while modelling the interactions between the two. Despite the simplicity of the model, some results can be deduced from the model. The propositions given above show what happens in a “classic” multicultural setting where there are more than one set of values coexisting in a population. It turns out that this setting is harder to achieve than may be thought and one such setting- where we have different sets of values but with everyone playing the same physical strategy, simply cannot exist as a stable stochastic equilibrium. Some other plausible settings such as those put forward in propositions 2 and 4 show that even populations with different strategies being played only coexist in particular ways when we have different values.

However, the most straightforward, and arguably the most interesting result, is that in the vast bulk of possible cases, the population will simply converge on one set of values. Indeed, even situations where there is a mixed equilibrium of physical strategies usually happen when the population only holds one set of values. It follows that, given the assumptions of the model, one would expect a population to converge on value uniformity with everyone sharing the same set of values. This undermines the notion that a multicultural society can survive without external “help”.

This last result means that, if one is to set up a multicultural society with values surviving long-term beside each other then one needs to construct institutions or rules that *prevent* the operation of the learning mechanisms in the model. There are various ways in which this could be done. One radical method is through a process of “ghettoisation” where a population with one set of values is isolated from another majority population by physical location or by a set of rules. In this case learning would only take place within the isolated population and so there would be no transmission of values from the majority population. Even if parts of the population are not ghettoised then institutions such as firms may be able to influence a person’s *work* values by partially isolating them from other influences.

Another possibility is that laws are passed and fines imposed that change the payoffs in the game. By changing the material payoffs one may hope to at least affect the utilities involved[13]. In doing so, one could change the utilities within each representation to such an extent that they change to a different category that enables them to coexist. A similar process could occur if laws are passed that increase psychic costs to such a level that representations similarly change. Alternatively, laws may block off certain representations being transmitted- an example would be the laws against racial hatred.

An important way in which values may be sustained is through education since this provides a “short-cut” in learning values. This may be used in sustaining a multicultural society by training part of the population in one set of values. An example of this may be the creation of religious schools where religious values are incorporated into the curriculum. Another example may be where the school curriculum aims to replace all values within a population wholesale with a fixed set of values. An example of this may be the inculcation of scientific values such as objectivity and accuracy in school, replacing more primitive superstitious ideas.

[13] However one should not believe that this is a simple matter. There have been increasing questions asked about the clash between intrinsic and extrinsic motivation. (c.f. Frey & Jegen 2001). There seems to be a tendency, for example, for intrinsic motivation to be wiped out by extrinsic incentives. This would mean that the use of fines may have perverse effects on behaviour.

It follows that institutions and rules will form the values of the populations that are subject to them simply by intruding in the evolutionary process by which values come to be common within a population. Conversely, the dismantling of these rules and institutions could have a disruptive effect on the survival of certain values within a population. As we have seen, in the vast bulk of cases, there will be a tendency towards the assimilation of diverse values to one particular set.

The question of how norms can be transferred from one situation to another can also be solved by the current model. This can be done by realising that in novel situations one has to decide how to interpret the situation and the actions of others in that situation. This means (as was stated in section 9) that one needs to find a suitable representation. The best way is to find a similar situation in one's past history and to use the representation for that situation. In other words, one needs to use a common psychological process, that of analogy, to find a good fit. The mystery of why one would, for example, leave a tip in a restaurant that one has no intention of visiting again becomes plain. One leaves a tip because one has made an analogy with a representation of other restaurant situations and this restaurant is a good "fit" with that representation. The representation's attached utilities then act as reasons for leaving a tip[14].

14. Conclusion

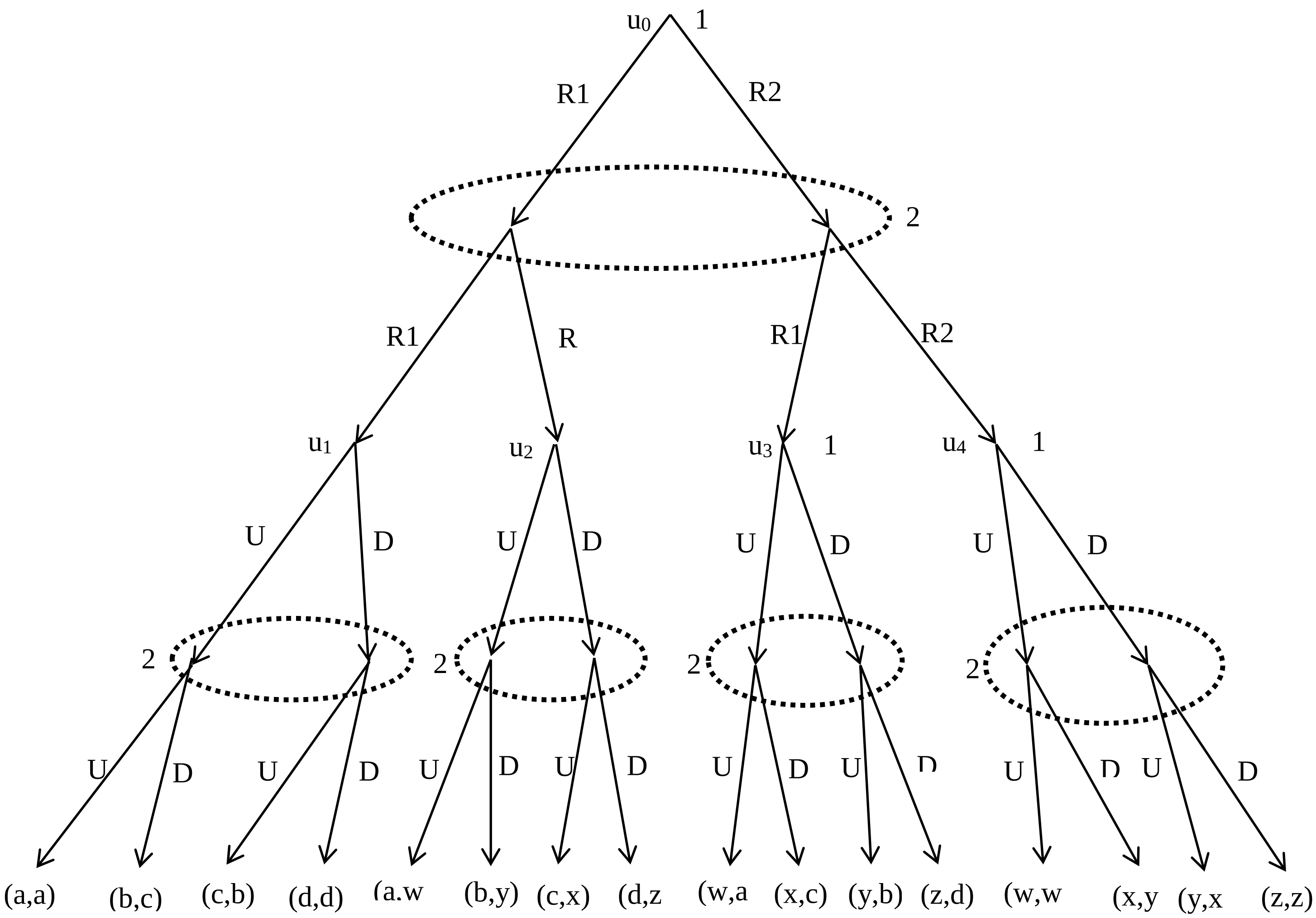
It is argued in this paper that the current models of values in economics are insufficient to explain many of the stylised facts that exist about values and, as a result, cannot say much about the role of values in the economy and society. Part of the problem is that economists have ignored two ideas that have been widely used within the psychological and philosophical literature for a long period of time, namely the idea of a mental action and the idea of a representation.

When we take account of these factors then we can build a simple model that is surprisingly productive in terms of predictions and can shed light on current debates in multiculturalism adding structure and objectivity to a debate that is often highly politicised. Furthermore, we can use the model to solve more abstract difficulties relating to the creation of social norms and the question of the transmission of norms.

[14] Note that this is different from the explanation offered by Sugden (2005). For Sugden the reason for giving a tip is the result of legitimate expectations, which have normative content- one thinks that the giving of a tip is ethically right and so one does it in all situations. Here, normativity and the transmission of normative behaviour are separate. The transmission is the result of analogies made with representations of previous similar situations while normativity is the result of valuation.

However, it should be noted that this model is still quite simple. A more realistic version would take account of the role of rules and institutions in value formation. Furthermore, the model is restricted to a two strategy base game. There is no reason in principle why the model should not be expanded outwards to include more physical strategies and more representations. However, this may mean that some of the more specific conclusions in the paper may have to be modified.

Figure 1



u0 1
R1 R2
2
R1 R R1 R2
u1 u2 u3 1 u4 1
U D U D U D U D
2 2 2 2
U D U D U D U D U D U D U D U D
(a,a) (b,c) (c,b) (d,d) (a.w (b,y) (c,x) (d,z (w,a (x,c) (y,b) (z,d) (w,w (x,y (y,x (z,z)

<u>Bibliography</u>

Asch S.E. (1940) “Studies in the principles of judgments and attitudes:II. Determination of judgments by group and by ego standards.” Journal of Social Psychology 12, p. 433-465

Bacharach M., (2006), “Beyond Individual Choice ; Team and Frame in Game Theory” ed. By Gold N., Sugden R., Princeton University Press

Becker, Gordon M., De Groot M., Marschak J. (1963) “Stochastic Models of Choice Behavior” Behavioral Science, 8 p.41

Berner A. & Putterman L. (eds) (1998) “Economics Values and Organization” Cambridge University Press Cambridge

Binmore K (1988) “Modelling Rational Players II” Economics and Philosophy vol. 4 p. 9-55

Bisin A. & Verdier T.(2001) “The Economics of Cultural Transmission and the Dynamics of Preferences” Journal of Economic Theory 97, 298-319

Bowles S. & Gintis H. (2011) “A Cooperative Species: Human Reciprocity and its evolution” Princeton University Press Princeton

Cavalli-Sforza, L.L., Menozzi P., & Piazza A. (1981) Cultural Transmission and Evolution: A Quantitative Approach Princeton, Princeton University Press

Cressman R. (2003) “Evolutionary Dynamics and Extensive Form Games” MIT Press Cambridge Massachusetts

Davidson D. (1963) “Actions, Reasons and Causes” Symposium on Action The Journal of Philosophy 60 685-700

Dawkins R. (1976) The Selfish Gene Oxford Oxford University Press

Denzau A.T. & North D.C. (1994) “Shared mental models: Ideologies and institutions. Kyklos, Vol. 47, Issue 1

Fehr E. & Schmidt K.M. (1999) “A Theory of Fairness, Competition and Cooperation” Quarterly Journal of Economics 114 p. 817-868

Frank R. (1988) “Passions within Reason: The Strategic Role of Emotions” W.W. Norton & Company New York

Frey B.S. & Jegen R. (2001) “Motivation Crowding Theory” Journal of Economic Surveys 15 p. 589-611

Fudenberg D. and Harris C. (1992) “Evolutionary Dynamics with Aggregate Shocks” Journal of Economic Theory 57 420-441

Geach P.(1957) “Mental Acts: Their Content and their Objects” London Routledge & Kegan Paul.

Gentner D., Holyoak K.J. & Kokinov B.N. (2001) The Analogical Mind: Perspectives from Cognitive Science MIT Press Cambridge Massachusetts

Henrich J. & Boyd R. (2002) “On Modelling Cognition and Culture: Why Cultural Evolution does not Require Replication of Representations” Journal of Cognition and Culture 87-112

Henrich J., Boyd R., Bowles S., Camerer C., Fehr E., and Gintis H. (2004) “Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies” Oxford Oxford University Press

Inglehart R., Basañez M., and Moreno A. (1998) “Human values and beliefs :political, religious, sexual and economic norms in 43 societies. Findings from the 1990-1993 World Values Survey “ Ann Arbor, Michigan. : University of Michigan Press

Jong, E. de, (2009) “Culture and economics :on values, economics and international business” London Routledge,.

Lewis D. (1969) “Convention: A philosophical Study” Wiley- Blackwell Oxford

Locke J. (1998) “An Essay Concerning Human Understanding” (first published 1689) Wordsworth editions Limited Hertfordshire

Mandler G. “Approaches to a Psychology of Value” in Hechter M., Nadel L. & Michod R.E (eds) (1993) “The Origin of Values” Walter de Gruyter Inc. New York

Mansbridge J. “Starting with nothing: on the impossibility of grounding norms in self-interest” in Berner A. & Putterman L. (eds) (1998) “Economics Values and Organization” Cambridge University Press Cambridge

Parekh B. (2006) “Rethinking Multiculturalism” (2nd Edition) Palgrave Macmillan Basingstoke Hampshire.

Proust J. (2001) “A Plea for Mental Acts” Synthese 129 105-128

Raz J. (2003) “The Practice of Value” The Berkeley Tanner Lectures Oxford University Press Oxford

Rokeach, M. (1973). The Nature of Human Values. New York: The Free Press.

Ross L. & Nisbett R.E. (1991) “The Person and the Situation: Perspectives of Social Psychology” McGraw- Hill Boston Massachussetts

Seligman C. & Katz A. “The Dynamics of Value Systems” in Seligman C., Olson J.M. & Zanna M. (eds) “The Psychology of Values” The Ontario symposium on Personality and Social Psychology vol. 8 (1996) Lawrence Erlbaum New Jersey.

Sen A. (1997) “Maximization and the Act of Choice” Econometrica 65 p. 745-779

Sen A. (1980) “Description as Choice” Oxford Economic Papers 32 p. 353-369

Sen A. (1971) “Choice Functions and Revealed Preference” Review of Economic Studies 38 p. 307-17

Simon H. (1959) “Theories of Decision-making in Economics and Behavioral Science” American Economic Review vol. 49 3 p. 253-283

Scharleman J., Eckel C. Kacelnik A. & Wilson R. (2001) “The Value of a Smile: Game Theory with a Human Face” Journal of Economic Psychology 22 p. 617-640.

Sperber D.(1996) “Explaining Culture: a naturalistic approach” Blackwell Oxford

Sugden, R. (2005) “The economics of rights, co-operation, and welfare” New York : Palgrave Macmillan.

Tabellini G. (2008) “The Scope of Cooperation: Values and Incentives” Quarterly Journal of Economics p. 905-950

Trevor-Roper H. “The Invention of tradition: The Highland tradition of Scotland” in Hobsbawm E. , Ranger T. (1983) (eds) The Invention of Tradition Cambridge University Press,

Neumann, J von and Morgenstern, O., (1953) Theory of Games and Economic Behavior. Princeton, NJ. Princeton University Press,.

Weibull J.W. (1995) “Evolutionary Game Theory” MIT Press Cambridge Massachussets.

Young H.P. (1993) “The Evolution of Conventions” Econometrica 61 p.57-84

Appendix 1

Demonstration of decomposition of differential equations:

Strategies in the extensive form take the form:

(X,t,v)

Where X is the choice at node u_0, t is the choice at nodes u_1 or u_3 and v is the choice at nodes u_2 or u_4.

There are eight possible strategies that have eight probabilities q_i (i=1..8).

	R_1UU	R_1UD	R_1DU	R_1DD	R_2UU	R_2UD	R_2DU	R_2DD
R_1UU	a,a	a,a	b,c	b,c	a,w	b,y	a,w	b,y
R_1UD	a,a	a,a	b,c	b,c	c,x	d,z	c,x	d,z
R_1DU	c,b	c,b	d,d	d,d	a,w	b,y	a,w	b,y
R_1DD	c,b	c,b	d,d	d,d	c,x	d,z	c,x	d,z
R_2UU	w,a	w,a	x,c	x,c	w,w	x,y	w,w	x,y
R_2UD	w,a	w,a	x,c	x,c	y,x	z,z	y,x	z,z
R_2DU	y,b	y,b	z,d	z,d	w,w	x,y	w,w	x,y
R_2DD	y,b	y,b	z,d	z,d	y,x	z,z	y,x	z,z

For ease of calculation assume that:

$$X_{ij} = \sum_{k=1}^{4} (q_{j\times k} \times v_{j\times k})$$

For i=1…8 and j=1 or 2 where v is a payoff.

This gives an abbreviated table:

	R_1UU	R_1UD	R_1DU	R_1DD	R_2UU	R_2UD	R_2DU	R_2DD
R_1	X_{11}	X_{21}	X_{31}	X_{41}	X_{51}	X_{61}	X_{71}	X_{81}
R_2	X_{12}	X_{22}	X_{32}	X_{42}	X_{52}	X_{62}	X_{72}	X_{82}

We can also use the following abbreviations:

$X^*_{11}=X_{11}=X_{21}$	$X^*_{12}=X_{31}=X_{41}$
$X^*_{21}=X_{12}=X_{22}$	$X^*_{22}=X_{32}=X_{42}$
$X^*_{31}=X_{51}=X_{71}$	$X^*_{32}=X_{61}=X_{81}$
$X^*_{41}=X_{52}=X_{72}$	$X^*_{51}=X_{62}=X_{82}$

To analyse the truncated game:

Assume that each payoff is stochastic in the sense that the population judgement of an individual payoff may vary. This variation is uniform across payoffs in the same representation but varies between the two representations. If k=1..4 and j=1..8 then

When R_1 is chosen by the population then the payoff change is $u_{kj}dt+\sigma_{R1}dW_1$

When R_2 is chosen by the population then the payoff change is $u_{kj}dt+\sigma_{R2}dW_2$

Where u represents the non-stochastic component of the payoffs

Each q_i moves according to the replicator dynamic, so:

$$\begin{aligned} dq_1 = q_1(\big(&(X_{11}+X_{12})dt+\sigma_{R1}dW_1\big) \\ &-\Big(q_1\big((X_{11}+X_{12})dt+\sigma_{R1}dW_1\big)+q_2\big((X_{21}+X_{22})dt+\sigma_{R1}dW_1\big) \\ &+q_3\big((X_{31}+X_{32})dt+\sigma_{R1}dW_1\big)+q_4\big((X_{41}+X_{42})dt+\sigma_{R1}dW_1\big) \\ &+q_5\big((X_{51}+X_{52})dt+\sigma_{R2}dW_2\big)+q_6\big((X_{61}+X_{62})dt+\sigma_{R2}dW_2\big) \\ &+q_7\big((X_{71}+X_{72})dt+\sigma_{R2}dW_2\big)+q_8\big((X_{81}+X_{82})dt+\sigma_{R2}dW_2\big)\Big)) \end{aligned}$$

Similarly for other strategies

If :

$$P_{R1} = q_1+q_2+q_3+q_4$$

Then:

$$dP_{R1} = dq_1+dq_2+dq_3+dq_4$$

$$\begin{aligned} dP_{R1} = \Big(&\big(q_1(X_{11}+X_{12})+q_2(X_{21}+X_{22})+q_3(X_{31}+X_{32})+q_4(X_{41}+X_{42})\big)dt \\ &+\ P_{R1}\sigma_{R1}dW_1\Big) \\ &-P_{R1}\Big(\big(q_1(X_{11}+X_{12})+q_2(X_{21}+X_{22})+q_3(X_{31}+X_{32}) \\ &+q_4(X_{41}+X_{42})\big)dt+P_{R1}\sigma_{R1}dW_1\Big)+((q_5(X_{51}+X_{52})+q_6(X_{61}+X_{62}) \\ &+q_7(X_{71}+X_{72})+q_8(X_{81}+X_{82})+(1-P_{R1})\sigma_{R2}dW_2) \end{aligned}$$

Assuming that:

$$\bar{n}_1 = \frac{(q_1 + q_2)X_{11}^* + (q_3 + q_4)X_{12}^*}{(q_1 + q_2 + q_3 + q_4)^2}$$

$$\bar{n}_2 = \frac{(q_1 + q_2)X_{21}^* + (q_3 + q_4)X_{22}^*}{(q_1 + q_2 + q_3 + q_4)(q_5 + q_6 + q_7 + q_8)}$$

$$\bar{n}_3 = \frac{(q_5 + q_7)X_{31}^* + (q_6 + q_8)X_{32}^*}{(q_1 + q_2 + q_3 + q_4)(q_5 + q_6 + q_7 + q_8)}$$

$$\bar{n}_4 = \frac{(q_5 + q_7)X_{41}^* + (q_6 + q_8)X_{42}^*}{(q_5 + q_6 + q_7 + q_8)^2}$$

These are the average payoffs if (R_1, R_1), (R_1,R_2),(R_2,R_1) and (R_2,R_2) are played respectively. If the subgames starting from u_1, u_2, u_3 and u_4 are in equilibrium then $\bar{n}_1, \bar{n}_2, \bar{n}_3$ *and* $\bar{n}_4$ are the equilibrium values for the respective subgames.

Substituting:

$$\begin{aligned} dP_{R1} = P_{R1}\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2)dt + \sigma_{R1}dW_1\big) \\ - ((P_{R1}(P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2)dt + P_{R1}\sigma_{R1}dW_1) \\ + (1 - P_{R1})(P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4)dt + (1 - P_{R1})\sigma_{R2}dW_2)) \end{aligned}$$

$$= P_{R1}(1 - P_{R1})\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4)\big)dt + \sigma_{R1}dW_1 - \sigma_{R2}dW_2)$$

Suppose :

$$\sigma = \sqrt{(\sigma_{R1}^2 + \sigma_{R2}^2)}$$

And: $W = \frac{(\sigma_{R1}W_1 - \sigma_{R2}W_2)}{\sigma}$

Therefore:

$$dP_{R1} = P_{R1}(1 - P_{R1})\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4)\big)dt + P_{R1}(1 - P_{R1})\sigma dW$$

For individual subgames:

Suppose v=1,2

Payoffs from playing strategy e_i and e_j under representation R_v are:

$$\pi(q\backslash e_i) = (X_{i1} + X_{i2})dt + \sigma_{Rv}dW_v$$

$$\pi(q\backslash e_j) = (X_{j1} + X_{j2})dt + \sigma_{Rv}dW_v$$

Suppose that v is fixed at v=v*. q_i^u is defined as the probability of taking physical action i within subgame u:

$$q_i^u = \frac{\sum_{i=1}^{2} q_i}{\sum_{j=1}^{4} q_j}$$

Using differentials:

$$dq_i^u = \frac{\sum_{i=1}^{2} dq_i \sum_{j=1}^{4} q_j - \sum_{i=1}^{2} q_i \sum_{j=1}^{4} dq_j}{\left(\sum_{j=1}^{4} q_j\right)^2}$$

Supposing:

$$dq_i = q_i(\pi(q\backslash e_i) - \pi(q))$$

And:

$$dq_j = q_j(\pi(q\backslash e_j) - \pi(q))$$

Gives:

$$dq_i^u = \frac{\sum_{i=1}^{2} \sum_{j=1}^{4} q_i q_j (\pi(q\backslash e_i) - \pi(q\backslash e_j))}{\left(\sum_{j=1}^{4} q_j\right)^2}$$

Substituting:

$$dq_i^u = \frac{\sum_{i=1}^{2}\sum_{j=1}^{4} q_i q_j \left(\left((X_{i1}+X_{i2})dt + \sigma_{Rv^*} dW_{v^*}\right) - \left(\left(X_{j1}+X_{j2}\right)dt + \sigma_{Rv^*} dW_{v^*}\right)\right)}{\left(\sum_{j=1}^{4} q_j\right)^2}$$

Cancelling:

$$dq_i^u = \frac{\sum_{i=1}^{2}\sum_{j=1}^{4} q_i q_j \left((X_{i1}+X_{i2}) - \left(X_{j1}+X_{j2}\right)\right)}{\left(\sum_{j=1}^{4} q_j\right)^2} dt$$

This can be reconfigured as a conventional differential equation:

$$\dot{q}_i^u = \frac{\sum_{i=1}^{2}\sum_{j=1}^{4} q_i q_j \left((X_{i1}+X_{i2}) - \left(X_{j1}+X_{j2}\right)\right)}{\left(\sum_{j=1}^{4} q_j\right)^2}$$

(The rest of the proof follows Cressman (2003 p. 185))

Looking at the numerator of the above equation, the expectations in the two payoff terms can be decomposed over endpoints m∈M with γ(.) being the probability of reaching a given endpoint. Hence the j payoff term can be decomposed into:

$$\sum_{m\in M}\sum_{i=1}^{2}\sum_{j=1}^{4} q_i q_j \gamma(m, q\backslash e_j)\pi(m)$$

For m that do *not* follow u, write e_j as $e^u_{j1\backslash j2}$ where j_1 refers to choices at those information sets of player 1 that are disjoint from u and j_2 refers to others then, on the Wright manifold:

$$\sum_{i=1}^{2}\sum_{j=1}^{4} q_i q_j \gamma(m, q\backslash e_j)\pi(m)$$

$$= \sum_{i=1}^{2}\sum_{j_1=1}^{2}\sum_{j_2=1}^{2} q^u_{i\backslash j_2} q_{j_1\backslash 1} \gamma(m, q\backslash e^u_{i\backslash j_2})\pi(m)$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{4} q^u_{i\backslash i} q^u_{j\backslash j} \gamma(m, q\backslash e^u_{i\backslash i})\pi(m)$$

On pathways disjoint from u this means that the difference between payoff terms in the numerator of the difference equation becomes 0. The whole equation becomes:

$$\dot{q}_i^u = \frac{\sum_{i=1}^{2}\sum_{j=1}^{4}\sum_{m\ follows\ u} q_i q_j (\gamma(m, q\backslash e_{i\backslash i}^u) - \gamma(m, q\backslash e_{j\backslash j}^u))\pi(m)}{\left(\sum_{j=1}^{4} q_j\right)^2}$$

$$= \frac{\sum_{m\ follows\ u}(\sum_{j=1}^{4} q_j \sum_{i=1}^{2} q_i \gamma(m, q\backslash e_{i\backslash i}^u) - \sum_{i=1}^{2} q_i \sum_{j=1}^{4} q_j \gamma(m, q\backslash e_{j\backslash j}^u))\pi(m)}{\left(\sum_{j=1}^{4} q_i\right)^2}$$

$$= \frac{\sum_{m\ follows\ u} \pi(m)(q_1^u \gamma(m, q\backslash e_i) - q_1^u \sum_{j=1}^{4} q_j \gamma(m, q\backslash e_j))}{\sum_{j=1}^{4} q_j}$$

$$= \frac{\sum_{m\ follows\ u} q_1^u (\gamma(m, q\backslash e_i)\pi(m) - \sum_{j=1}^{4} q_j \gamma(m, q\backslash e_j)\pi(m))}{\sum_{j=1}^{4} q_j}$$

$$= q_1^u (K^u(q\backslash e_i)\pi(q^u\backslash e_1^u) - K^u(q\backslash e_j)\pi(q^u))$$

Where K^u is the probability that node u is reached.

$$= K^u(q\backslash e_i) q_1^u (\pi^u(q^u\backslash e_1^u) - \pi^u(q^u))$$

It follows that the game has one stochastic truncated game equation:

$$dP_{R1} = P_{R1}(1 - P_{R1})\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4)\big)dt + P_{R1}(1 - P_{R1})\sigma dW$$

And four deterministic subgame equations:

Subgame 1:

$$\dot{q}_U^{u_1} = P_{R1}^{u_0} q_U^{u_1} \left[q_U^{u_1} a + (1 - q_U^{u_1})b - \left[q_U^{u_1}(q_U^{u_1} a + (1 - q_U^{u_1})b) + (1 - q_U^{u_1})(q_U^{u_1} c + (1 - q_U^{u_1})d\right]\right]$$

Subgame 2:

$$\dot{q}_U^{u_2} = (1 - P_{R1}^{u_0}) q_U^{u_2} \Big[q_U^{u_3} a + \big(1 - q_U^{u_3}\big) b - \big[q_U^{u_2} \big(q_U^{u_3} a + \big(1 - q_U^{u_3}\big) b \big) + (1 - q_U^{u_2}) (q_U^{u_3} c + \big(1 - q_U^{u_3}\big) d \big] \Big]$$

Subgame 3:

$$\dot{q}_U^{u_3} = P_{R1}^{u_0} q_U^{u_3} \Big[q_U^{u_2} w + \big(1 - q_U^{u_2}\big) x - \big[q_U^{u_3} \big(q_U^{u_2} w + \big(1 - q_U^{u_2}\big) x \big) + (1 - q_U^{u_3}) (q_U^{u_2} y + \big(1 - q_U^{u_2}\big) z \big] \Big]$$

Subgame 4:

$$\dot{q}_U^{u_4} = (1 - P_{R1}^{u_0}) q_U^{u_4} \Big[q_U^{u_4} w + \big(1 - q_U^{u_4}\big) x - \big[q_U^{u_4} \big(q_U^{u_4} w + \big(1 - q_U^{u_4}\big) x \big) + (1 - q_U^{u_4}) (q_U^{u_4} y + \big(1 - q_U^{u_4}\big) z \big] \Big]$$

Proof of Proposition 1:

The proof of this proposition closely follows that of Fudenburg and Harris (1992)

Sokhorod (1989) shows that:

If $I_1=\infty$ and $I_2 \neq \infty$ then converges to R_1

If $I_1 \neq \infty$ and $I_2=\infty$ then converges to R_2

If $I_1 \neq \infty$ and $I_2 \neq \infty$ then converges to R_1 with probability $I_1/(I_1+I_2)$ and R_2 with probability $I_2/(I_1+I_2)$

If $I_1=\infty$ and $I_2 \neq \infty$ then system oscillates forever. This can be refined using $M(P_{R1})$. If $M(P_{R1})$ is finite then there is a unique ergodic distribution:

$$\frac{M(P_{R1})}{\int_0^1 M(P_{R1}) dP_{R1}}$$

The distribution converges on this as $t \to \infty$.

Substituting from the stochastic equation:

$$I_1 = \int_0^{P_{R1}(0)} exp\left[- \int_\alpha^q \frac{2\big((P_{R1}\bar{n}_1 + (1 - P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1 - P_{R1})\bar{n}_4) \big)}{P_{R1}(1 - P_{R1})\sigma^2} dP_{R1} \right] dP_{R1}$$

$$I_2 = \int_{P_{R1}(0)}^{1} exp\left[-\int_{\alpha}^{q} \frac{2\big((P_{R1}\bar{n}_1 + (1-P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1-P_{R1})\bar{n}_4)\big)}{P_{R1}(1-P_{R1})\sigma^2} dP_{R1}\right] dP_{R1}$$

$$M(P_{R1}) = \frac{2}{\sigma^2} exp\left[\int_{\alpha}^{q} \frac{2\big((P_{R1}\bar{n}_1 + (1-P_{R1})\bar{n}_2) - (P_{R1}\bar{n}_3 + (1-P_{R1})\bar{n}_4)\big)}{P_{R1}(1-P_{R1})\sigma^2} dP_{R1}\right]$$

The inner integrals for I_1 and I_2 are identical so they can be integrated and placed in the outer integral as follows:

$$I_1 = \int_0^{P_{R1}(0)} \left(\frac{P_{R1}}{\alpha}\right)^{\frac{2(\bar{n}_4-\bar{n}_2)}{\sigma^2}} \left(\frac{1-P_{R1}}{1-\alpha}\right)^{\frac{2(\bar{n}_1-\bar{n}_3)}{\sigma^2}} dP_{R1}$$

$$I_2 = \int_{P_{R1}(0)}^{1} \left(\frac{P_{R1}}{\alpha}\right)^{\frac{2(\bar{n}_4-\bar{n}_2)}{\sigma^2}} \left(\frac{1-P_{R1}}{1-\alpha}\right)^{\frac{2(\bar{n}_1-\bar{n}_3)}{\sigma^2}} dP_{R1}$$

Using similar reasoning $M(P_{R1})$ can also be integrated:

$$M(P_{R1}) = \frac{2}{\sigma^2}\left[\left(\frac{P_{R1}}{\alpha}\right)^{\frac{2(\bar{n}_2-\bar{n}_4)}{\sigma^2}} \left(\frac{1-P_{R1}}{1-\alpha}\right)^{\frac{2(\bar{n}_3-\bar{n}_1)}{\sigma^2}}\right]$$

I_1 is integrated between 0 and $P_{R1}(0)$ so, given the limits, it is an improper integral if the exponent on the (P_{R1}/α) term is negative. It is divergent if the exponent is less than or equal to -1 and convergent otherwise.

I_2 is integrated between $P_{R1}(0)$ and 1 so a similar argument can be made with the ((1-P_{R1})/(1-α)) term.

A similar argument can be made for the integral:

$$\int_0^1 M(P_{R1}) dP_{R1}$$

Hence: I_1 is finite iff :

$$\frac{2(\bar{n}_4 - \bar{n}_2)}{\sigma^2} > -1$$

Or: $$\bar{n}_2 - \bar{n}_4 < \sigma^2/2$$

Likewise I_2 is finite iff:

$$\frac{2(\bar{n}_1 - \bar{n}_3)}{\sigma^2} > -1$$

Or: $$\bar{n}_3 - \bar{n}_1 < \sigma^2/2$$

The integral of $M(P_{R1})$ is finite iff:

$$\frac{2(\bar{n}_2 - \bar{n}_4)}{\sigma^2} > -1 \quad \text{and} \quad \frac{2(\bar{n}_3 - \bar{n}_1)}{\sigma^2} > -1$$

Or: $\bar{n}_4 - \bar{n}_2 < \sigma^2/2$ and $\bar{n}_1 - \bar{n}_3 < \sigma^2/2$

If $\varphi = \int_0^1 M(P_{R1})dP_{R1}$ then:

$$\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}} = \frac{1}{\varphi}\left(\frac{P_{R1}}{\alpha}\right)^{\frac{2(\bar{n}_2-\bar{n}_4)}{\sigma^2}}\left(\frac{1-P_{R1}}{1-\alpha}\right)^{\frac{2(\bar{n}_3-\bar{n}_1)}{\sigma^2}}$$

Proof of proposition 2:

First note that the (1v4) and (4v1) cases are identical so we only need to prove one to prove the other. We will prove the (4v1) case.

Assume subgame monotonicity and generic payoffs with (4v1) as the category combination.

For the symmetric subgame following u_1, use equation (2*):

$$\dot{q}_U^{u_1} = P_{R1}q_U^{u_1}(1 - q_U^{u_1})\left[q_U^{u_1}(a - c) + \left(1 - q_U^{u_1}\right)(b - d)\right]$$

Assume that this is a category 4 game (i.e. a<c and b<d) so it must be the case that $q_U^{u_1} \to 0$ as t→∞.

For the symmetric subgame following u_4, use equation (5*)

$$\dot{q}_U^{u_4} = (1 - P_{R1})q_U^{u_4}(1 - q_U^{u_4})\left[q_U^{u_4}(w - y) + \left(1 - q_U^{u_4}\right)(x - z)\right]$$

Assume that this is a category 1 game (i.e. w>y and x>z) so it must be the case that $q_U^{u_4} \to 1$ as $t \to \infty$

For the linked asymmetric subgame, assuming the same payoffs and use equations (3*) and (4*)

$$\dot{q}_U^{u_2} = (1 - P_{R1})q_U^{u_2}(1 - q_U^{u_2})\left[q_U^{u_3}(a - c) + \left(1 - q_U^{u_3}\right)(b - d)\right]$$

$$\dot{q}_U^{u_3} = P_{R1}q_U^{u_3}(1 - q_U^{u_3})\left[q_U^{u_2}(w - y) + \left(1 - q_U^{u_2}\right)(x - z)\right]$$

This results in $q_U^{u_2} \to 0$ as $t \to \infty$ and $q_U^{u_3} \to 1$ as $t \to \infty$. This results in the following truncated game payoff matrix only showing deterministic payoffs:

	R_1	R_2
R_1	d,d	c,x
R_2	x,c	w,w

Assume that the population has a stochastic equilibrium with stable probability distribution $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$ as defined by Proposition 1. This means that d-x< $\sigma^2/2$ and w-c<$\sigma^2/2$. By inspection of the payoff matrix, one can see that those members of the population that select Representation 1 will also only play D (since d and c only ever appear as utility outcomes of playing D) while those members of the population that select Representation 2 will also only play U (since x and w only ever appear as utility outcomes of playing U).

Now assume that we have only one strategy played in one representation and the other strategy only played in the second representation. We also assume that we have a mixed equilibrium in the truncated game with a probability distribution $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$.

None of the outcomes in the truncated game can be category 3 because the proposition excludes mixed action solutions by assumption. (2v3) is excluded anyway by assumption.

In each symmetric subgame there are a maximum of only two pure strategies equilibria that the population can converge on. This means there are two possible cases in a given symmetric subgame where the population converges to an equilibrium.

The symmetric subgame following u_1 is reached by playing R_1 while the symmetric subgame following u_4 is reached by playing R_2. Suppose that in the symmetric subgame following u_1, $q_U^{u_1} \to 0$ as t→∞. It follows that in the symmetric subgame following u_4, $q_U^{u_4} \to 1$ as t→∞ because everyone choosing R_2 must choose the opposite strategy by assumption. For there to be just one strategy played in one representation then for role 1 (playing R_1) in the linked asymmetric subgame $q_U^{u_2} \to 0$ as t→∞ as this is consistent in physical strategy with the subgame following u_1. For role 2 (playing R_2) $q_U^{u_3} \to 1$ as t→∞ as this is consistent in physical strategy with the subgame following u_4. Excluding category 3, this pattern of preferences can only occur with the category combination (4v1).

A similar argument establishes (1v4).■

Proof of proposition 3:

Proof by contradiction.

By proposition 1 if the population has $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$ as an equilibrium stochastic equilibrium then we are in a mixed strategy case for representations.

Assume that such an equilibrium exists. Also assume that we have generic payoffs. Fix an arbitrary strategy (since it is a symmetric game, we will choose U).

In the symmetric subgames: since category 3 is excluded and since we have assumed pure strategies so $q_U^{u_1} \to 1$ and $q_U^{u_4} \to 1$ as t→∞.

Assuming one has the same payoffs in the linked asymmetric subgame then $q_U^{u_2} \to 1$ and $q_U^{u_3} \to 1$ as t→∞.

This means that the truncated game matrix, with deterministic payoffs only, is as follows:

	R_1	R_2
R_1	a,a	a,w
R_2	w,a	w,w

Given generic payoffs, there can only be two possible preference orderings: a-w> $\sigma^2/2$ or a-w< $\sigma^2/2$. In the first case $P_{R1} \to 1$ as t→∞ with probability 1 while in the second case $P_{R1} \to 0$ as t→∞ with probability 1. This contradicts the assumption of an equilibrium stochastic equilibrium with distribution $\frac{M(P_{R1})}{\int_0^1 M(P_{R1})dP_{R1}}$. A similar argument can be made for setting a physical strategy of D. Hence the proposition is true. ■

Proof of proposition 4:

Firstly note that it is not possible to have a situation where there are any mixed strategies in the asymmetric subgames because (2v3) and (3v2) are excluded by assumption.

Given this, there is only one way in which the proposition can hold i.e. if the symmetric subgames have mixed strategies equilibria. This means that $q_U^{u_1} \rightarrow \frac{d-b}{d-b+a-c}$ and $q_U^{u_4} \rightarrow \frac{z-x}{z-x+w-y}$ as t→∞. These ratios must both be positive and less than 1 so either the numerator and denominator are positive or they are both negative. To be less than 1, the differences (z-x), (w-y), (d-b) and (a-c) must either be all positive or all negative. This suggests that they can only exist as equilibria under categories 2 or 3.

Examining equations 2* and 5*:

$$\dot{q}_U^{u_1} = P_{R1} q_U^{u_1}(1 - q_U^{u_1})\left[q_U^{u_1}(a-c) + \left(1 - q_U^{u_1}\right)(b-d)\right]$$
$$\dot{q}_U^{u_4} = (1 - P_{R1}) q_U^{u_4}(1 - q_U^{u_4})\left[q_U^{u_4}(w-y) + \left(1 - q_U^{u_4}\right)(x-z)\right]$$

It can be seen that they only both converge on their interior equilibria if a<c, b>d,w<y and x>z. If this is the case then the category combination must be (3v3)■